Chin. Ann. Math.
27B(3), 2006, 243–262
DOI: 10.1007/s11401-005-0176-y

Chinese Annals of Mathematics, Series B

Talagrand's T_2-Transportation Inequality and Log-Sobolev Inequality for Dissipative SPDEs and Applications to Reaction-Diffusion Equations***

Liming WU*   Zhengliang ZHANG**

Abstract We establish Talagrand's T_2-transportation inequalities for infinite dimensional dissipative diffusions with sharp constants, through Galerkin type's approximations and the known results in the finite dimensional case. Furthermore in the additive noise case we prove also logarithmic Sobolev inequalities with sharp constants. Applications to Reaction-Diffusion equations are provided.

Keywords Stochastic partial differential equations (SPDEs), Logarithmic Sobolev inequality, Talagrand's transportation inequality, Poincaré inequality
2000 MR Subject Classification 60H15, 37L40, 35K57, 35R60

1 Introduction

The purpose of this paper is to study the large time behavior of the Stochastic Differential Equation (SDE in short) in a Hilbert space H,

$$\begin{cases} dX(t) = (AX(t) + F(X(t)))dt + B(X(t))dW(t), \quad t > 0, \\ X(0) = x \in H, \end{cases}$$

satisfying the dissipative condition (see Section 2 for details). The known study is usually realized from the dynamical or analytical point of view, e.g., establishing the exponential convergence in some functions space of transition semigroup or log-Sobolev inequality for its invariant measure, see [7] and [5] and references therein. In this work we shall adopt the probabilistic point of view, i.e., establishing the Talagrand T_2-transportation inequality and the log-Sobolev inequality on the path space $C([0,T],H)$ (or $L^2([0,T],H)$) for the law of $X(\cdot)$ with respect to (w.r.t. in short) the $L^2([0,T],H)$-metric. Those inequalities, according to Ledoux [13] et al., imply the sharp concentration inequalities for

$$\mathbb{P}\Big(\frac{1}{T}\int_0^T V(X(s))ds \in \cdot\Big),$$

a central object in the probabilistic understanding of the ergodic behavior of X. Let us first recall the transportation inequality.

Manuscript received May 13, 2005.
*Laboratoire de Mathématiques Appliquées, CNRS-UMR 6620, Université Blaise Pascal, 63177 Aubière, France; School of Mathematics and Statistics, Wuhan University, Wuhan 430072, China.
E-mail: Li-Ming.Wu@math.univ-bpclermont.fr
**School of Mathematics and Statistics, Wuhan University, Wuhan 430072, China.
E-mail: zzl731205@yahoo.com.cn
***Project supported by the Yangtze Scholarship Program.

Let (E, d) be a metric space equipped with σ-field $\mathcal{B}$ such that $d(\cdot, \cdot)$ is $\mathcal{B}\times\mathcal{B}$-measurable. Given $p \in [1, +\infty]$ and two probability measures μ and ν on E, the Wasserstein distance is defined by

$$W_p(\mu, \nu) = \inf \Big(\iint d^p(x, y) d\pi(x, y) \Big)^{1/p}, \tag{1.1}$$

where the infimum is taken over all probability measures π on the product space $E \times E$ with marginal distributions μ and ν (saying couplings of (μ, ν)).

The relative entropy of ν w.r.t. μ, denoted by $H(\nu/\mu)$, is given by

$$H(\nu/\mu) = \begin{cases} \int \log \frac{d\nu}{d\mu} d\nu, & \text{if } \ \nu \ll \mu, \\ +\infty, & \text{otherwise.} \end{cases} \tag{1.2}$$

We say that the probability measure μ satisfies the L^p-transportation inequality on (E, d) if there exists a constant $C \geq 0$ such that for any probability measure ν,

$$W_p(\mu, \nu) \leq \sqrt{2CH(\nu/\mu)}\,. \tag{1.3}$$

To be short we write $\mu \in T_p(C)$ for this relation. The cases "$p = 1$" and "$p = 2$" are particularly interesting. That $T_1(C)$ is related to the phenomenon of measure concentration was emphasized by K. Marton [14, 15], Bobkov and Götze [3] and amply explored by Ledoux [13]. Recently H. Djellout, A. Guillin and the first named author [4] proved the following criterion: $\mu \in T_1(C)$ iff for some $\delta > 0$,

$$\iint_{E^2} e^{\delta d^2(x,y)} \mu(dx)\mu(dy) < +\infty.$$

The $T_2(C)$, stronger than $T_1(C)$, was first established by Talagrand [18] for the Gaussian measure, and it has been brought into relation with the log-Sobolev inequality, Poincaré inequality, inf-convolution, Hamilton-Jacobi's equations etc. by Otto-Villani [16] and Bobkov-Gentil-Ledoux [2]. The work of Talagrand on the Gaussian measure has been generalized on an abstract Wiener space by Feyel and Ustunel [9, 10]. F. Y. Wang [20] obtained the $T_2(C)$ w.r.t. the L^2-metric on path spaces over Riemannian manifolds by tensorization. H. Djellout, A. Guillin and the first named author [4] studied the T_2-transportation inequality w.r.t. the L^2-metric for finite dimensional diffusions by means of Girsanov transformation and for general dependent sequences by using the coupling method of Marton [14].

More recently F. Y. Wang [21] establishes that inequality w.r.t. the intrinsic metric on the path spaces, and S. Fang and J. Shao [8] prove the $T_2(C)$ w.r.t. the Cameron-Martin metric on path and loop groups. And in [22], we obtain the $T_2(C)$ for finite dimensional diffusions w.r.t. a uniform metric.

Let us recall now the log-Sobolev inequality. Let μ be a probability measure on some separable Hilbert space H. We say that μ satisfies the log-Sobolev inequality with the constant $C \geq 0$, denoted by $\mu \in \mathrm{logS}(C)$, if

$$\mu(f^2 \log f^2) - \mu(f^2) \log \mu(f^2) \leq 2C\mu(|\nabla f|_H^2), \quad \forall\, f \in C_b^1(H). \tag{1.4}$$

Here $C_b^1(H)$ denotes the space of all real continuously differentiable functions f on H such that f and the gradient ∇f are bounded, and $\mu(f) := \int f d\mu$.

The log-Sobolev inequality is one of most important tools in the infinite dimensional analysis and it is widely and deeply studied in the past thirty years since its introduction by Gross [12] (see [13] for a systematic treatment of the subject). It is known by the work of Bobkov-Gentil-Ledoux [2] that $\log S(C) \implies T_2(C)$ on $H = \mathbb{R}^d$ and their argument shows that the same implication holds true on a general separable Hilbert space H.

The main purpose of this work is to generalize the results in H. Djellout, A. Guillin and the first named author [4] about finite dimensional diffusions to dissipative stochastic partial differential equations (SPDE in short). Furthermore in the additive noise case, we obtain the log-Sobolev inequality which is stronger than the $T_2(C)$. In that case our results generalize a recent work by Da Prato, Debussche and Goldys [5] in two respects: first our results cover the case of reaction-diffusion in higher dimension (only one-dimensional case was treated in [5]); second, our results are not only for the marginal law and the invariant measure, but also for the law of the whole path of the underlying SPDE, which lead to sharp concentration inequalities describing the ergodic behavior of the diffusion.

This paper is organized as follows. The main results are presented in the next section, where we give T_2-transportation inequalities for Lipschitzian system with general noise and log-Sobolev inequalities for dissipative system with additive noise. In Section 3, applications to stochastic reaction-diffusion equations are provided. In Section 4, we introduce some known results in the finite dimensional case. Section 5 and Section 6 are devoted to the proofs of the main results. Our approach consists in approximating the solution of infinite dimensional SPDE by finite dimensional diffusions, which might have independent interests.

2 Main Results

2.1 Notations

Let H be a separable Hilbert space endowed with the norm $|\cdot|$ induced by its inner product, U another separable Hilbert space. The Banach space of all linear and bounded (resp. and Hilbert-Schmidt) operators from U into H is denoted by $L(U \to H)$ (resp. $L_2(U \to H)$), with the operator norm $\|\cdot\|$ (resp. Hilbert-Schmidt norm $\|\cdot\|_2$). Let $C([0,T], H)$ be the space of all continuous functions from $[0,T]$ into H and

$$L^2([0,T],H) = \Big\{ f : [0,T] \to H \text{ measurable; } |f|_2^2 := \int_0^T |f(t)|^2 dt < +\infty \Big\}.$$

The L^2-metric on $L^2([0,T],H) \supset C([0,T],H)$ is defined as follows,

$$d_2(\gamma_1, \gamma_2) := |\gamma_1 - \gamma_2|_2 = \sqrt{\int_0^T |\gamma_1(t) - \gamma_2(t)|^2 dt}, \quad \forall\, \gamma_1, \gamma_2 \in L^2([0,T],H).$$

Consider the Stochastic Differential Equation (SDE in short) in the Hilbert space H,

$$\begin{cases} dX(t) = (AX(t) + F(X(t)))dt + B(X(t))dW(t), \quad t > 0, \\ X(0) = x \in H, \end{cases} \tag{2.1}$$

where (W_t) is a cylindrical Brownian motion in the Hilbert space U with the identity covariance operator I_U, defined on some well filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\ge 0}, \mathbb{P})$ ("well filtered" means that $(\mathcal{F}_t)_{t\geq 0}$ satisfies the usual condition); A is the generator of some C_0-semigroup on H, $B : H \to L(U \to H)$; and $F : \mathbb{D}(F)(\subset H) \to H$.

For short, the Stochastic Differential Equation with coefficients A, F, B such as (2.1) is written as SDE (A, F, B). Let us introduce

Hypothesis 2.1 (i) *A is the infinitesimal generator of some C_0 semigroup $(S(t))_{t\geq 0}$ on H;*

(ii) *For any $u \in U$, $x \to B(x)u$ is continuous from H to H, and for any $t > 0$ and $x \in H$, $S(t)B(x) \in L_2(U \to H)$, and there is some nonnegative locally square-integrable function $K(t)$ on $\mathbb{R}^+$ such that for all $x, y \in H$,*

$$\begin{aligned} &\|S(t)B(x)\|_2 \leq K(t)(1 + |x|), \\ &\|S(t)B(x) - S(t)B(y)\|_2 \leq K(t)|x - y|; \end{aligned} \tag{2.2}$$

(iii) $\sup_{x\in H} \|B(x)\|_{U\to H} \leq M < \infty.$

Hypothesis 2.2 (Dissipativity) *There is some $\delta > 0$ such that for all $x, y \in \mathbb{D}(A) \bigcap \mathbb{D}(F)$,*

$$\langle x - y, A(x - y) + F(x) - F(y)\rangle + \frac{1}{2}\|B(x) - B(y)\|_2^2 \leq -\delta|x - y|^2. \tag{2.3}$$

Here $\mathbb{D}(A)$ (resp. $\mathbb{D}(F)$) is the domain of definition of A (resp. F) in H.

Notice that under Hypothesis 2.2, $B(x) - B(y) \in L_2(U \to H)$ but $B(x)$ alone may be not Hilbert-Schimidt.

2.2 Lipschitzian non-linearality case

We first consider the case where the non-linear term F is Lipschitzian.

(L) $\mathbb{D}(F) = H$ and $F : H \to H$ is Lipschitzian, i.e., there exists some $L > 0$ such that

$$|F(x) - F(y)| \leq L|x - y|, \quad \forall\, x, y \in H.$$

Theorem 2.1 *Assume Hypotheses* 2.1 *and* 2.2 *and* (L). *Suppose furthermore that*

Hypothesis 2.3 *there is a sequence of finite dimensional orthogonal projections $(\Pi_n)_{n\in\mathbb{N}}$ such that*

(i) *the range H_n of Π_n is contained in $\mathbb{D}(A) \bigcap \mathbb{D}(F)$;*

(ii) *for any $x \in \mathbb{D}(A)$, there exists $x_n \in H_n$ such that $|x_n - x| + |Ax_n - Ax| \to 0$;*

(iii) *For all n, $e^{t\Pi_n A\Pi_n}\Pi_n B(\Pi_n x)$ satisfies* (2.2) *with the same $K(t)$ (in place of $S(t)B(x)$) and for each $t > 0$, $x \in H$,*

$$\|e^{t\Pi_n A\Pi_n}\Pi_n B(\Pi_n x) - S(t)B(x)\|_2 \to 0. \tag{2.4}$$

Then the SDE (2.1) *has a unique L^2-mild solution $X(t) = X(t, x)$, i.e., $X(t)$ is progressively measurable, $\sup_{t\leq T} \mathbb{E}|X(t)|^2 < +\infty$ for every $T > 0$ and satisfies for each $t \geq 0$ fixed,*

$$X(t) = S(t)x + \int_0^t S(t - s)[F(X(s))ds + B(X(s))dW_s], \quad \mathbb{P}\text{- } a.s.$$

Moreover this process $X(t, x)$ possesses the following properties:

(a) *For all different initial points $x, y \in H$,*

$$\mathbb{E}|X(t,x) - X(t,y)|^2 \le e^{-2\delta t}|x-y|^2, \quad \forall\, t \ge 0. \tag{2.5}$$

In particular $X(\,\cdot\,)$ has a unique invariant probability measure μ such that for any initial measure ν on H,

$$W_2(\nu P_t, \mu) \le e^{-\delta t} W_2(\nu, \mu), \quad \forall t \ge 0, \tag{2.6}$$

where $P_t(x, \cdot\,) := \mathbb{P}(X(t,x) \in \,\cdot\,)$, $W_2(\nu,\mu)$ is the L^2-Wasserstein distance between ν and μ w.r.t. the $|\cdot|$-metric of H.

(b) *the probability distribution $\mathbb{P}_x$ (on $L^2([0,T],H)$) of the mild solution $X(\,\cdot\,,x)$ starting from x of* SDE (A,F,B) (2.1) *satisfies the Talagrand transportation inequality $T_2(C)$ on $L^2([0,T],H)$ w.r.t. the L^2-metric d_2 for all $x \in H$ and $T > 0$, where the constant C is given by*

$$C := \frac{M^2}{\delta^2}.$$

Furthermore $P_T(x, \cdot\,) \in T_2(\frac{M^2}{2\delta})$ on $(H, |\cdot|)$, as well as the unique invariant probability measure μ of (P_t).

(c) *When $U = H$ and $B^{-1}(x)$ exists and satisfies*

$$\widetilde{M} := \sup_{x\in H} \|B^{-1}(x)\| = \sup_{x,z\in H;|z|=1} |B^{-1}(x)z| < +\infty,$$

the following Poincaré inequality holds

$$\mathrm{Var}_\mu(f) := \mu(f^2) - \mu(f)^2 \le \frac{M^2\widetilde{M}^2}{2\delta}\mu(|B\nabla f|^2), \quad \forall\, f \in C_b^1(H). \tag{2.7}$$

Remark 2.1 The first claim about the existence and the uniqueness of L^2-mild solution holds without Hypothesis 2.3 (see [7, Theorem 5.3.1]). Furthermore, if there exists some $\alpha \in (0, \frac{1}{2})$ such that for $\forall\, T > 0$,

$$\int_0^T t^{-2\alpha} K(t)^2 dt < +\infty,$$

where $K(t)$ is given in Hypothesis 2.1(ii), then $X(\,\cdot\,)$ allows a continuous version (see [7, Theorem 5.3.1]).

Remark 2.2 Assume Hypothesis 2.1. One can find a sequence of finite dimensional orthogonal projections (Π_n) satisfying Hypothesis 2.3 in each of the following situations (the details are left to the reader):

(1) $x \to B(x)$ is globally Lipschitzian from H to $L_2(U \to H)$;

(2) $S(t)$ is symmetric and compact on H (by the spectral decomposition) for each $t > 0$.

Remark 2.3 All inequalities in parts (a), (b), (c) in the theorem above are sharp, as seen for the Ornstein-Uhlenbeck process: $F = 0$, $U = H$ and $B(x) = I_U$ and $S(t)$ is symmetric such that $\|S(t)\|_2$ is locally square integrable on $[0, +\infty)$. See [4] for details.

Remark 2.4 That $\mu \in T_2(\frac{M^2}{2\delta})$ in part (b) implies the following Gaussian integrability:

$$\int_H e^{\lambda|x|^2} d\mu < +\infty, \quad \forall\, \lambda \in \Big(0, \frac{\delta}{M^2}\Big).$$

This fact together with (2.6) yields $W_2(P_T(x, \cdot\,), \mu) \to 0$ as T goes to infinity.

2.3 The dissipative and additive noise case

When the non-linearity of F becomes stronger, we have to assume that $B(x) = B$ (constant covariance matrix), i.e., the case of additive noise (as in the general theory of SPDEs, [6, 7]). The new framework is as follows.

Assume that K is a reflexive Banach space, densely and continuously embedded into H. We introduce the following hypotheses (see [7, p.80]).

Hypothesis 2.4 (i) *There exist $\eta_1,\ \eta_2 \in \mathbb{R}$ such that the operators $A - \eta_1$ and $F - \eta_2$ are m-dissipative on H and $\delta := -(\eta_1 + \eta_2) > 0$ (that is stronger than Hypothesis 2.2);*

(ii) *for some $\tilde{\eta}_1,\ \tilde{\eta}_2 \in \mathbb{R}$, the parts on K of $A - \tilde{\eta}_1$ and $F - \tilde{\eta}_2$ are m-dissipative on K;*

(iii) *$\mathbb{D}(F) \supset K$ and F maps bounded sets in K into bounded sets of H.*

Hypothesis 2.5 *The process*

$$W_A(t) = \int_0^t S(t-s)BdW(s), \quad t \geq 0,$$

allows a continuous version in H, takes values in the domain $\mathbb{D}(F_K)$ of the part of F in K, and for any $T > 0$,

$$\sup_{t\in[0,T]} (\|W_A(t)\|_K + \|F(W_A(t))\|_K) < +\infty, \quad \mathbb{P}\text{-}\ a.s.$$

Under Hypotheses 2.4 and 2.5, we know that (see [7, Theorem 5.5.8])

(i) for each $x \in K$, the SDE (2.1) has a unique continuous mild solution $X(\,\cdot\,, x)$;

(ii) for each $x \in H$, the SDE (2.1) has a unique continuous generalized mild solution $X(\,\cdot\,, x)$, i.e., there exists a sequence (x_n) in K and mild solutions $X(\,\cdot\,, x_n)$ such that $X(\,\cdot\,, x_n) \to X(\,\cdot\,, x)$ uniformly over bounded time intervals in probability.

Theorem 2.2 *Assume that Hypotheses 2.1 and 2.3–2.5 are fulfilled. For every $x \in H$, let $\mathbb{P}_x$ be the law of the unique generalized mild solution $X(\cdot, x)$ of the SDE (2.1) (a probability measure on $C([0,T],H) \subset L^2([0,T],H)$), and $P_t(x,dy) := \mathbb{P}(X(t,x) \in dy)$ the transition probability kernel. Then*

(a) *For all $x, y \in H$ and $t \geq 0$,*

$$|X(t,x) - X(t,y)| \leq e^{-\delta t}|x-y|, \quad a.s.$$

In particular P_t has a unique invariant measure μ such that for any $p \in [1, +\infty]$, any initial probability measure ν, and any $t \geq 0$,

$$W_p(\nu P_t, \mu) \leq e^{-\delta t} W_p(\nu, \mu).$$

(b) (T_2-Transportation Inequality) *$\mathbb{P}_x \in T_2(C)$ on $C([0,T],H)$ w.r.t. the L^2-metric d_2 for all $x \in H$ and $T > 0$, where the constant C is given by*

$$C := \frac{\|B\|^2}{\delta^2}. \tag{2.8}$$

Moreover $P_T(x,\,\cdot\,) \in T_2\big(\frac{\|B\|^2}{2\delta}\big)$ on H, as well as the unique invariant probability measure μ of (P_t).

(c) (Log-Sobolev Inequality) *The law $\mathbb{P}_x$ satisfies the* $\operatorname{logS}(C)$ *with* $C = \frac{\|B\|^2}{\delta^2}$ *on* $L^2([0,T], H)$*, i.e., for all* $F \in C_b^1(L^2([0,T],H))$,

$$\mathbb{E}F^2(X_{[0,T]}) \log \frac{F^2(X_{[0,T]})}{\mathbb{E}F^2(X_{[0,T]})} \leq 2\frac{\|B\|^2}{\delta^2}\mathbb{E}|\nabla F|_2^2(X_{[0,T]}), \tag{2.9}$$

where ∇ *is the gradient on the Hilbert space* $L^2([0,T],H)$*. Moreover, for every* $T > 0$,

$$P_T(f^2 \log f^2) - P_T f^2 \log P_T f^2 \leq \frac{\|B\|^2}{\delta} P_T(|\nabla f|_H^2), \quad \forall\, f \in C_b^1(H) \tag{2.10}$$

on H*, and* $\mu \in \operatorname{logS}(C)$ *with* $C = \frac{\|B\|^2}{2\delta}$.

(d) *In particular, when* $U = H$ *and* B^{-1} *exists and is bounded,*

$$\mu(f^2 \log f^2) - \mu(f^2) \log \mu(f^2) \leq \frac{\|B\|^2 \|B^{-1}\|^2}{\delta} \mu(|B\nabla f|_H^2), \quad \forall\, f \in C_b^1(H). \tag{2.11}$$

Remark 2.5 When $U = H$, $B = I_H$, $F = 0$ and A is symmetric, all $T_2(C)$ and $\operatorname{logS}(C)$ in the theorem above become sharp.

2.4 Several consequences of the T_2-transportation inequality

As shown in [4, 22], many interesting consequences can be derived from the $T_2(C)$ of $\mathbb{P}_x$ on $L^2([0,T],H)$ above. For instance (for the detailed proof, see [4]),

Corollary 2.1 *Under the assumptions of Theorem* 2.1 *or Theorem* 2.2*, let* $C := M^2\delta^{-2}$*. We have for any* $T > 0$,

(a) *The following Poincaré inequality holds for any* $F \in C_b^1(L^2([0,T],dt;H))$,

$$\operatorname{Var}_{\mathbb{P}_x}(F) \leq C \int_{L^2([0,T],H)} |\nabla F(\gamma)|_2^2 d\mathbb{P}_x(\gamma), \tag{2.12}$$

where $\operatorname{Var}_{\mathbb{P}_x}(F)$ *is the variance of* F *under law* $\mathbb{P}_x$*, and* $\nabla F(\gamma) \in L^2([0,T],H)$ *is the gradient of* F *at* γ *in* $L^2([0,T],H)$.

(b) (Inequality of Tsirel'son Type) *For any non-empty subset* G *in* $L^2([0,T],H)$ *such that* $Z(\gamma) := \sup_{h\in G}\langle \gamma, h\rangle \in L^1(\mathbb{P}_x)$ *where* $\langle h_1, h_2\rangle := \int_0^T \langle h_1(t), h_2(t)\rangle_H dt$ *is the inner product in* $L^2([0,T],H)$*, we have*

$$\int \exp\Big(\frac{1}{C}\sup_{h\in G}\Big[\langle \gamma, h\rangle - \frac{|h|_2^2}{2}\Big]\Big) d\mathbb{P}_x \leq \exp\Big(\frac{1}{C}\mathbb{E}^{\mathbb{P}_x} Z\Big). \tag{2.13}$$

(c) (Inequality of Hoeffding Type) *For any Lipschitz function* $V : H \to \mathbb{R}$ *such that its Lipschitzian coefficient* $\|V\|_{\mathrm{Lip}} \leq \alpha$*, we have for all* $r > 0$ *and* $T > 0$,

$$\mathbb{P}\Big(\frac{1}{T}\int_0^T V(X(t,x))dt - \mathbb{E}\frac{1}{T}\int_0^T V(X(t,x))dt > r\Big) \leq \exp\Big(-\frac{CTr^2}{2\alpha^2}\Big). \tag{2.14}$$

See Ledoux [13] for further important consequences of the log-Sobolev inequalities in Theorem 2.2.

3 Applications to Stochastic Reaction-Diffusion Equations

In this section, we will study the following stochastic reaction-diffusion equation on a bounded open domain $\mathcal{O}$ in $\mathbb{R}^N$:

$$\begin{cases} dX(t,\xi) = (\Delta X(t,\xi) + f(X(t,\xi)))dt + \sqrt{Q}\, dW(t,\xi), \quad t > 0, \\ X(0,\xi) = x(\xi), \quad \forall\, \xi \in \mathcal{O}, \\ X(t,\xi) = X(t,\xi) = 0, \quad \xi \in \partial\mathcal{O}, \end{cases} \tag{3.1}$$

where $U = H = L^2(\mathcal{O}) := L^2(\mathcal{O}, dx)$, $\xi \in \mathcal{O}$, Δ is the Laplacian operator with the Dirichlet boundary condition, i.e., $\mathbb{D}(\Delta) = H^2(\mathcal{O}) \bigcap H_0^1(\mathcal{O})$, $f : \mathbb{R} \to \mathbb{R}$, and $Q : L^2(\mathcal{O}) \to L^2(\mathcal{O})$ is a bounded nonnegative definite symmetric operator, and $W(t,\xi)$ is a cylindrical Wiener process with the covariance I_H, defined on some well filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$.

By a mild solution of (3.1), we mean an adapted process $X(t)$ such that $t \to X(t)$ is continuous from $\mathbb{R}^+$ to $L^2(\Omega, \mathbb{P}; H)$ where $H = L^2(\mathcal{O})$ and for all $t \geq 0$,

$$X(t) = S(t)x + \int_0^t S(t-s)F(X(s))ds + \int_0^t S(t-s)\sqrt{Q}\, dW_s,$$

where $S(t)$ is the Dirichlet semigroup generated by $(\Delta, \mathbb{D}(\Delta) = H^2(\mathcal{O}) \bigcap H_0^1(\mathcal{O}))$, $F(x)(\xi) := f(x(\xi))$. A generalized mild solution $X(t,x)$ of (3.1) means that there is a sequence of mild solutions $X(t,x_n), n \in \mathbb{N}$ (with initial conditions x_n) such that $X(t,x_n) \to X(t,x)$ in probability for every $t \geq 0$.

Notice that the boundary condition at $\partial\mathcal{O}$ is assured formally by the Dirichlet boundary condition of Δ.

3.1 The one-dimensional case

Proposition 3.1 *Let $N = 1$ and $\mathcal{O} = (0,1)$. Assume that f satisfies*

Hypothesis 3.1 *$f(x) + \lambda x$ is continuous and non-increasing for some $\lambda \in \mathbb{R}$, and for some $s \geq 1$ and $c_0 > 0$,*

$$|f(x)| \leq c_0(1 + |x|^s), \quad x \in \mathbb{R}.$$

Assume moreover that

$$\delta := \pi^2 + \lambda > 0. \tag{3.2}$$

Then for every initial condition $x \in L^2(0,1)$, the SDE (3.1) *has a unique generalized mild solution $X(\,\cdot\,, x)$, whose law is denoted by $\mathbb{P}_x$. Furthermore all parts* (a), (b), (c) *and* (d) *in Theorem* 2.2 *hold true with $\|B\| = \|\sqrt{Q}\,\| = \sqrt{\|Q\|}$, the constant δ given above and $\|B^{-1}\| = \sqrt{\|Q^{-1}\|}$.*

Remark 3.1 If f is derivable on $\mathbb{R}$ with f' upper bounded and is of polynomial growth at infinity, Hypothesis 3.1 is satisfied with $\lambda = -\sup_{z\in\mathbb{R}} f'(z)$. For example, if

$$f(x) = -\alpha x^{2n+1} + \sum_{k=0}^{2n} b_k x^{2n-k}, \quad x \in \mathbb{R},\ \alpha > 0,$$

then Hypothesis 3.1 is satisfied with $s = 2n+1$, and when $\alpha > 0$ is sufficiently large so that $\delta = \pi^2 + \lambda > 0$.

Remark 3.2 The log-Sobolev inequality (2.11) in Theorem 2.2(d) for this one-dimensional model was proved by G. Da Prato, A. Debussche and B. Goldys [5], with a completely different approach from ours. But other conclusions in this proposition are new.

Proof of Proposition 3.1 It is enough to verify all assumptions of Theorem 2.2 with $H = L^2(0,1)$ and $K = L^{2s}(0,1)$.

(1) We begin with Hypothesis 2.1. Since $A = \frac{d^2}{d\xi^2}$ with $\mathbb{D}(A) = H^2(0,1) \bigcap H_0^1(0,1)$ is the infinitesimal generator of the Dirichlet semigroup $S(t)$ on $H = L^2(0,1)$, and $S(t) \in L_2(H),\ t > 0$, indeed,

$$\|S(t)\|_2^2 = \sum_{n=1}^{\infty} e^{-2n^2\pi^2 t}$$

which is equivalent to $\frac{c}{\sqrt{t}}$ for some $c > 0$ as $t \to 0$. Hence (2.2) is satisfied with $K(t) = \|S(t)\|_2$.

(2) (Verification of Hypothesis 2.3) Since Hypothesis 2.1 is satisfied and $S(t)$ is symmetric and Hilbert-Schmidt for every $t > 0$, Hypothesis 2.3 is automatically satisfied by Remark 2.2.

(3) (Verification of Hypothesis 2.4) At first $A+\pi^2$ is m-dissipative on $L^2(0,1)$ (well-known), i.e., $A - \eta_1$ is dissipative with $\eta_1 = -\pi^2$. Moreover for every $p \in [1, +\infty)$, A is m-dissipative (for the Dirichlet semigroup $S(t)$ generated by A is a C_0-semigroup of contractions on $L^p(0,1)$).

Let us show that $F(x)(\xi) := f(x(\xi))$ satisfies : $F + \lambda$ is m-dissipative on $L^p(0,1)$ for every $p \geq 1$. Indeed for all $s > t \in \mathbb{R}$, $(s-t)[f(s) - f(t) + \lambda(s-t)] \leq 0$. So we have for all $\beta > 0$,

$$|s - t - \beta[f(s) - f(t) + \lambda(s-t)]| \geq |s-t|.$$

By symmetry it continues to hold true for $s < t$. Thus for all $x, y \in L^p(0,1)$ such that $F(x), F(y) \in L^p(0,1)$, we have

$$\|x - y - \beta[F(x) - F(y) + \lambda(x-y)]\|_{L^p(0,1)} \geq \|x-y\|_{L^p(0,1)},$$

that is, $F + \lambda$ is dissipative, and furthermore it is m-dissipative (the detail is left to the reader). Thus (i) and (ii) in Hypothesis 2.4 are verified.

Finally it is obvious that $F : K = L^{2s}(0,1) \to L^2(0,1) = H$ is continuous and bounded, then Hypothesis 2.4(iii) is satisfied.

(4) (Verification of Hypothesis 2.5) To this end (with $K = L^{2s}(0,1)$), it is enough to establish $\mathbb{E} \sup_{0\leq t\leq T} \|W_A(t)\|^p_{L^p(0,1)} < \infty$ for all $p \geq 2s$.

Since for any $\alpha \in (0, \frac{1}{4})$,

$$\int_0^T t^{-2\alpha} \|S(t)\sqrt{Q}\,\|_2^2 dt \leq \|Q\| \int_0^T t^{-2\alpha} \|S(t)\|_2^2 dt < +\infty, \quad \forall T > 0,$$

$W_A(t) := \int_0^t S(t-s)\sqrt{Q}\,dW_t$ has a continuous version in the domain $\mathbb{D}((-A)^\gamma)$ for any $\gamma \in (0, \alpha)$, by [7, Theorem 5.2.6]. That implies $W_A(t, \xi)$ has a continuous version in both $(t, \xi) \in \mathbb{R}^+ \times [0,1]$ for $\mathbb{D}((-A)^\gamma)$ is continuously embedded into $C[0,1]$. This is much stronger than what we require.

3.2 The multi-dimensional case

We now discuss the case where $N > 1$. In the multi-dimensional case, if $Q = I_H$ in (3.1), the Hypothesis 2.5 is not satisfied for $\int_0^T \|S(t)\|_2^2 dt = +\infty$ in general (its finiteness is a necessary condition for defining $\int_0^t S(t-s)dW(s)$). Indeed if $N=2$ and $\mathcal{O}=(0,\pi)^N$,

$$\|S(t)\|_2^2 \geq \sum_{n,m=1}^{\infty} e^{-2n^2t}e^{-2m^2t} = \Big(\sum_{n=1}^{\infty} e^{-2n^2t}\Big)^2 \sim t^{-1}.$$

In order to overcome this difficulty, we need to impose some conditions on the covariance matrix Q of the noise.

Proposition 3.2 *Let $N>1$ and $\mathcal{O}$ be a bounded open domain of $\mathbb{R}^N$. Assume that f satisfies Hypothesis* 3.1. *Suppose moreover that*

$$\delta := \lambda_0(\mathcal{O}) + \lambda > 0, \tag{3.3}$$

where $\lambda_0(\mathcal{O}) := \inf\{\int_{\mathcal{O}} |\nabla f|^2 dx;\ f \in H_0^1(\mathcal{O}), \int_{\mathcal{O}} f^2 dx = 1\}$ is the minimal eigenvalue of $-\Delta$ with the Dirichlet boundary condition in $\mathcal{O}$; and for some $\alpha\in(0,\frac{1}{2})$ and $T>0$,

$$\int_0^T t^{-2\alpha}\|S(t)\sqrt{Q}\|_2^2 dt = \int_0^T t^{-2\alpha}\mathrm{tr}(S(t)QS(t)^*)dt < +\infty. \tag{3.4}$$

Then for every initial condition $x\in L^2(\mathcal{O})$, the SDE (3.1) *has a unique generalized mild solution $X(\,\cdot\,,x)$, and all conclusions of Theorem* 2.2 *except part* (d) *hold true.*

Proof The proof is completely identical to that of Proposition 3.1, except (3.4), being automatically satisfied in the one-dimensional case, becomes now a condition.

Let us exhibit a concrete situation where condition (3.4) is satisfied. Since A^{-1} is compact and symmetric (for $\mathcal{O}$ is bounded), there is an orthonormal basis $(e_k)_{k\geq 0}$ of $L^2(\mathcal{O})$ such that

$$Ae_k = -\lambda_k(\mathcal{O})e_k, \quad \forall\, k\geq 0,$$

where the sequence of $(\lambda_k(\mathcal{O}))_{k\geq 0}$ increases to infinity.

The cylindrical Wiener process $W(t,\xi)$ on $H=L^2(\mathcal{O})$ can be always written as follows (formally):

$$W(t,\xi) = \sum_{k\geq 0} e_k(\xi)\beta_k(t), \quad t\geq 0,$$

where $\beta_k(\,\cdot\,),\ k\in\mathbb{N}$ are independent standard real-valued Brownian motion defined on $(\Omega,\mathcal{F},(\mathcal{F}_t),\mathbb{P})$.

Assume now that Q is determined by

$$Qe_k(\xi) = q_k e_k, \quad 0<q_k\leq c,\ \forall\, k\in\mathbb{N}. \tag{3.5}$$

Corollary 3.1 *Assume that Q is determined by* (3.5). *If for some $\gamma\in(0,1)$,*

$$\sum_{k\in\mathbb{N}} \frac{q_k}{\lambda_k(\mathcal{O})^{1-\gamma}} < +\infty, \tag{3.6}$$

then the condition (3.4) *in Proposition* 3.2 *is verified. In particular if moreover Hypothesis* 3.1 *and* (3.3) *are satisfied, then all conclusions of Proposition* 3.2 *hold true.*

Proof Note that for any $\alpha\in(0,\frac{1}{2})$ and $\lambda>0$,

$$\int_0^{+\infty} t^{-2\alpha}e^{-\lambda t}dt = \frac{c}{\lambda^{1-2\alpha}},$$

where $c = \int_0^{+\infty} t^{-2\alpha} e^{-t} dt$. Then for any $0 < \alpha < \frac{\gamma}{2}$, we have

$$\int_0^{+\infty} t^{-2\alpha} \|S(t)\sqrt{Q}\,\|_2^2 dt = \sum_{k\geq 0} \int_0^{+\infty} t^{-2\alpha} e^{-2\lambda_k(\mathcal{O})t} q_k dt \leq \sum_{k\geq 0} \frac{cq_k}{\lambda_k(\mathcal{O})^{1-2\alpha}} < +\infty.$$

So this corollary follows by Proposition 3.2.

4 Known Results in the Finite Dimensional Case

We begin with the finite dimensional case. Consider the diffusion process

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t, \quad X_0 = x, \tag{4.1}$$

where $b: \mathbb{R}^d \to \mathbb{R}^d$, and $\sigma : \mathbb{R}^d \to M_{d\times n}$ (the space of $d \times n$ matrices), and (B_t) is the standard Brownian Motion valued in $\mathbb{R}^n$ defined on some well filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$. Assume the following dissipative condition:

$$\langle x - y, b(x) - b(y)\rangle + \frac{1}{2}\mathrm{tr}(\sigma(x) - \sigma(y))(\sigma(x) - \sigma(y))^t \leq -\delta|x - y|^2, \quad \forall\, x, y \in \mathbb{R}^d, \tag{4.2}$$

where $\delta > 0$ and A^t denotes the transposition of matrix A.

Theorem 4.1 (See [4]) *Assume that b, σ are locally Lipschitzian and (4.2) holds, and $\|\sigma\| := \sup\{|\sigma(x)z|;\ x \in \mathbb{R}^d, |z| \leq 1\} < +\infty$. Let $\mathbb{P}_x$ be the law of the solution $(X(t, x))$ to the SDE (4.1) and $P_t(x, \cdot\,) := \mathbb{P}(X(t, x) \in \cdot\,)$. Then*

(a) *For all $t \geq 0, x, y \in \mathbb{R}^d$,*

$$\mathbb{E}|X(t, x) - X(t, y)|^2 \leq e^{-2\delta t}|x - y|^2. \tag{4.3}$$

In particular (P_t) has a unique invariant probability measure μ such that for any initial measure ν on H,

$$W_2(\nu P_t, \mu) \leq e^{-\delta t} W_2(\nu, \mu), \quad \forall\, t \geq 0,$$

where $W_2(\nu, \mu)$ is the L^2-Wasserstein distance between ν and μ w.r.t. the euclidean metric of $\mathbb{R}^d$.

(b) *$\mathbb{P}_x \in T_2(C)$ on $L^2([0, T], \mathbb{R}^d)$ w.r.t. the L^2-metric d_2 for all $x \in \mathbb{R}^d$ and $T > 0$, where the constant C is given by*

$$C := \frac{\|\sigma\|^2}{\delta^2}. \tag{4.4}$$

Moreover $P_T(x, \cdot\,) \in T_2\big(\frac{\|\sigma\|^2}{2\delta}\big)$ on $\mathbb{R}^d$, as well as the unique invariant probability measure μ of (P_t).

Corollary 4.1 *Under the conditions of Theorem 4.1, if $n = d$ and $\sigma^{-1}(x)$ exists and satisfies $\|\sigma^{-1}\| := \sup_{x\in H} \|\sigma^{-1}(x)\| = \sup_{x,z\in H; |z|=1} |B^{-1}(x)z| < +\infty$, then the following Poincaré inequality holds*

$$\mathrm{Var}_\mu(f) := \mu(f^2) - \mu(f)^2 \leq \frac{\|\sigma\|^2\|\sigma^{-1}\|^2}{2\delta}\mu(|\sigma\nabla f|^2), \quad \forall\, f \in C_b^1(\mathbb{R}^d);$$

as well as the following exponential convergence in $L^2(\mathbb{R}^d, \mu)$:

$$\mathrm{Var}_\mu(P_t f) \leq \exp\Big(-\frac{\|\sigma\|^2\|\sigma^{-1}\|^2}{\delta}t\Big)\mathrm{Var}_\mu(f), \quad \forall\, f \in L^2(\mathbb{R}^d, \mu),\ \forall\, t \geq 0.$$

Proof It is known (see [2]) that $\mu \in T_2(C)$ implies the following Poincaré inequality

$$\mathrm{Var}_\mu(f) \le C\mu(|\nabla f|^2), \quad \forall\, f \in C_b^1(\mathbb{R}^d).$$

Thus the Poincaré inequality in the corollary follows from Theorem 4.1(b). For the L^2-exponential convergence, by approximating b, σ by smooth ones, we can assume without loss of generality that $b, \sigma \in C^\infty$ moreover. In that case, since part (a) implies

$$|\nabla P_t f|^2 \le e^{-2\delta t} P_t(|\nabla f|^2),$$

the family $\mathcal{A} := \{f \in C_b^1(\mathbb{R}^d) \bigcap C^\infty(\mathbb{R}^d); \mathcal{L}f \text{ bounded}\}$ is stable by (P_t), where $\mathcal{L} = \frac{1}{2}\mathrm{tr}(\sigma\sigma^t\nabla^2 f)$ $+b(x)\cdot\nabla f$ is the generator of (P_t). In that smooth case the L^2-exponential convergence for $f \in \mathcal{A}$ is equivalent to the Poincaré inequality for $f \in \mathcal{A}$. Finally it is easy to extend the L^2-exponential convergence from $f \in \mathcal{A}$ to the whole space $f \in L^2(\mu)$.

When $\sigma(x) = \sigma$ is a constant matrix, we have the following stronger log-Sobolev inequalities.

Theorem 4.2 *Under the assumptions of Theorem* 4.1, *suppose moreover that* $\sigma(x) = \sigma$ *is a constant matrix. Then*

(a) *For all* $t \ge 0,\ x, y \in \mathbb{R}^d$,

$$|X(t,x) - X(t,y)| \le e^{-\delta t}|x - y|, \quad a.s. \tag{4.5}$$

In particular for all $p \in [1, +\infty]$, *initial measure* ν *and* $t \ge 0$,

$$W_p(\nu P_t, \mu) \le e^{-\delta t} W_p(\nu, \mu).$$

(b) (Log-Sobolev Inequality) *The following log-Sobolev inequality holds for the law* $\mathbb{P}_x$ *on* $L^2([0,T], \mathbb{R}^d)$*: for all* $F \in C_b^1(L^2([0,T],\mathbb{R}^d))$,

$$\mathbb{E}F^2(X_{[0,T]}) \log \frac{F^2(X_{[0,T]})}{\mathbb{E}F^2(X_{[0,T]})} \le 2\frac{\|\sigma\|^2}{\delta^2}\mathbb{E}|\nabla F|^2_{L^2([0,T],\mathbb{R}^d)}(X_{[0,T]}), \tag{4.6}$$

where ∇ *is the gradient on the Hilbert space* $L^2([0,T],\mathbb{R}^d)$.

(c) *For every* $T > 0$,

$$P_T\left(f^2 \log f^2\right) - P_T f^2 \log P_T f^2 \le \frac{\|\sigma\|^2}{\delta} P_T(|\nabla f|^2), \quad \forall\, f \in C_b^1(\mathbb{R}^d), \tag{4.7}$$

i.e., $P_T(x, \cdot) \in \mathrm{logS}\big(\frac{\|\sigma\|^2}{2\delta}\big)$ *on* $\mathbb{R}^d$. *And the unique invariant measure* μ *satisfies the same log-Sobolev inequality* (4.7) *instead of* P_T.

(d) *In particular, when* $n = d$ *and* σ^{-1} *exists,*

$$\mu(f^2 \log f^2) - \mu(f^2)\log\mu(f^2) \le 2\frac{\|\sigma\|^2\|\sigma^{-1}\|^2}{\delta}\mu(|\sigma\nabla f|^2), \quad \forall\, f \in C_b^1(\mathbb{R}^d), \tag{4.8}$$

and

$$\mathrm{Ent}_\mu(P_t f) \le \exp\Big(-\frac{\|\sigma\|^2\|\sigma^{-1}\|^2}{4\delta}t\Big)\mathrm{Ent}_\mu(f), \quad \forall\, 0 \le f \in L^1(\mu), \tag{4.9}$$

where $\mathrm{Ent}_\mu(f) = \mu(f\log f) - \mu(f)\log\mu(f) \in [0, +\infty]$, *is the entropy of* f, $0 \le f \in L^1(\mu)$.

Proof Since $\widehat{X}(t) := X(t,x) - X(t,y)$ satisfies

$$\frac{d}{dt}|\widehat{X}(t)|^2 = 2\langle \widehat{X}(t), b(X(t,x)) - b(X(t,y))\rangle \le -2\delta|\widehat{X}(t)|^2$$

by the dissipativity condition (4.2), we get (4.5) by Gronwall's inequality. The last exponential convergence in the L^p-Wasserstein distance follows easily from (4.5) (see [4] for details).

Part (b) was established by M. Gourcy and the first named author in [11].

For part (c), by approximation and Lemma 4.2, we may assume without loss of generality that $b \in C^\infty(\mathbb{R}^d)$. To prove (4.7), we follow the beautiful and classical argument of Bakry-Emery [1]. For any $f \in C_b^2(\mathbb{R}^d)$ which is uniformly positive (i.e., bounded from below by some constant $\varepsilon > 0$), we have for any $T > 0$ fixed and $t \in (0,T)$,

$$\begin{aligned}\frac{d}{dt}P_t(P_{T-t}f\log P_{T-t}f) &= P_t\mathcal{L}(P_{T-t}f\log P_{T-t}f) - P_t[(\log P_{T-t}f+1)\mathcal{L}P_{T-t}f]\\ &= \frac{1}{2}P_t\Big(\frac{1}{P_{T-t}f}|\sigma^*\nabla P_{T-t}f|^2\Big),\end{aligned}$$

so

$$\begin{aligned}P_T(f\log f) - P_Tf\log P_Tf &= \frac{1}{2}\int_0^T P_t\Big(\frac{1}{P_{T-t}f}|\sigma^*\nabla P_{T-t}f|^2\Big)dt\\ &\le \frac{\|\sigma\|^2}{2}\int_0^T P_t\Big(\frac{1}{P_{T-t}f}|\nabla P_{T-t}f|^2\Big)dt. \qquad (4.10)\end{aligned}$$

By (4.5), we have $|\nabla P_tf| \le e^{-\delta t}P_t(|\nabla f|)$. Thus we obtain by Cauchy-Schwartz inequality

$$\begin{aligned}P_t\Big(\frac{1}{P_{T-t}f}|\nabla P_{T-t}f|^2\Big) &\le e^{-2\delta(T-t)}P_t\Big(\frac{1}{P_{T-t}f}P_{T-t}(|\nabla f|)^2\Big)\\ &\le e^{-2\delta(T-t)}P_t\Big(\frac{1}{P_{T-t}f}P_{T-t}\Big(\frac{|\nabla f|^2}{f}\Big)P_{T-t}f\Big)\\ &= e^{-2\delta(T-t)}P_T\Big(\frac{|\nabla f|^2}{f}\Big).\end{aligned}$$

Substituting into (4.10), we obtain

$$P_T(f\log f) - P_Tf\log P_Tf \le \frac{\|\sigma\|^2}{4\delta}P_T\Big(\frac{|\nabla f|^2}{f}\Big).$$

Replacing f by f^2, we get (4.7) for all $f \in C_b^2(\mathbb{R}^d)$ which are uniformly positive. This can be easily extended to $f \in C_b^1(\mathbb{R}^d)$ by approximating $|f|$ by $f_n \in C_b^2(\mathbb{R}^d)$ which are uniformly positive.

Letting $T \to \infty$, as $P_T(x,\cdot) \to \mu$ weakly, we get (4.7) for μ by Lemma 4.2.

For part (d), note at first that (4.8) follows from (4.7). For the the exponential convergence (4.9) in entropy, let $\mathcal{A} := \{g \in C^\infty(\mathbb{R}^d)\bigcap C_b^1(\mathbb{R}^d); \exists\, \varepsilon > 0, 0 < \varepsilon \le g \le 1/\varepsilon, \mathcal{L}g \text{ bounded}\}$, where $\mathcal{L} = \frac{1}{2}\mathrm{tr}(\sigma\sigma^*\nabla^2 f) + b(x)\cdot\nabla f$ is the generator of P_t. Since $\mathcal{A}$ is stable by P_t (because $b \in C^\infty(\mathbb{R}^d)$ satisfies the dissipative condition (4.2)), one gets (4.9) for $f \in \mathcal{A}$ from (4.8) by differentiating $\mathrm{Ent}_\mu(P_tf)$ in t and Gronwall's inequality (well known). Finally the extension of (4.9) from $f \in \mathcal{A}$ to $0 \le f \in L^1(\mu)$ is easy.

We shall still need two general lemmas.

Lemma 4.1 (See [4]) *Let H be a separable Hilbert space and $(\mu_n)_{n\in\mathbb{N}}$ a sequence of probability measures on H which converges weakly to μ. Assume that for some constant $C \geq 0$, $\mu_n \in T_2(C)$ for all n w.r.t. the metric $d(x,y) = |x-y|$. Then $\mu \in T_2(C)$.*

Lemma 4.2 *Let H be a separable Hilbert space and $(\mu_n)_{n\in\mathbb{N}}$ a sequence of probability measures on H which converges weakly to μ. Assume that there exists some constant $C \geq 0$ such that $\mu_n \in \log\mathrm{S}(C)$ for all n, i.e.,*

$$\mu_n(f^2\log f^2) - \mu_n(f^2)\log\mu_n(f^2) \leq 2C\mu_n(|\nabla f|^2), \quad \forall\, f \in C_b^1(H). \tag{4.11}$$

Then μ satisfies the log-Sobolev inequality on H with the same constant $C \geq 0$.

Proof Its proof is obvious since $f^2, f^2\log f^2, |\nabla f|^2$ are all continuous and bounded on H for $f \in C_b^1(H)$.

5 Proof of Theorem 2.1: Galerkin's Approximation

Let (Π_n) be the sequence of finite dimensional projections on H, specified by Hypothesis 2.3. For each $n \geq 1$, define

$$A_n x := \Pi_n A \Pi_n x, \quad F_n(x) := \Pi_n F(\Pi_n x), \quad B_n(x) := \Pi_n B(\Pi_n x). \tag{5.1}$$

Since A_n is bounded (by the closed graph theorem), F_n is Lipschtzian and for any $x, y \in H_n = \mathrm{Ran}(\Pi_n)$,

$$\begin{aligned}\|B_n(x) - B_n(y)\|_2 &\leq \|e^{-tA_n}e^{tA_n}\Pi_n(B(x)-B(y))\|_2\\ &\leq \|e^{-tA_n}\|\cdot\|e^{tA_n}\Pi_n(B(x)-B(y))\|_2\\ &\leq \|e^{-tA_n}\|K(t)|x-y|,\end{aligned}$$

hence the coefficients in the SDE

$$dX_n(t) = (A_nX_n(t) + F_n(X_n(t)))dt + B_n(X_n(t))dW(t), \quad X_n(0) = \Pi_n x \tag{5.2}$$

are Lipschitzian and then (5.2) admits a unique solution $X_n(t)$.

Lemma 5.1 *Under the assumptions of Theorem* 2.1*, let $\mathbb{P}^n_{\Pi_n x}$ be the law of $X_n(\cdot)$, the solution to the SDE* (5.2)*. Then*

(a) *All conclusions of Theorem* 4.1 *hold true for $X_n(t)$, with $\mathbb{R}^d$ replaced by H, $\|\sigma\|$ replaced by M (in Hypothesis* 2.1*).*

(b) *When $B(x) = B$ is constant, all conclusions of Theorem* 4.2 *except part* (d) *hold true for $X_n(t)$, with $\mathbb{R}^d$ replaced by H, $\|\sigma\|$ replaced by M.*

Proof From (2.3), we have for all $x, y \in H_n$,

$$\begin{aligned}&\langle x-y, A_n(x-y) + F_n(x) - F_n(y)\rangle + \frac{1}{2}\|B_n(x) - B_n(y)\|_2^2\\ &\leq \langle x-y, A(x-y) + F(x) - F(y)\rangle + \frac{1}{2}\|B(x)-B(y)\|_2^2 \leq -\delta|x-y|^2.\end{aligned}$$

And

$$\|B_n\| := \sup_{x,z\in H_n, |z|\leq 1}|B_n(x)z| \leq \sup_{x,z\in H, |z|\leq 1}|B(x)z| \leq M.$$

By Theorem 4.1, part (a) holds on H_n, and thus on H.

The log-Sobolev inequalities in part (b) on H_n instead of H follow by Theorem 4.2. For general $F \in C_b^1(L^2([0,T],H))$ or $f \in C_b^1(H)$, since $\mathbb{P}((1-\Pi_n)X_n(t)=0,\ \forall t)=1$, the corresponding log-Sobolev inequalities follow from those for $F_n(\gamma) := F(\Pi_n\gamma(\cdot))$ or $f_n(x) := f(\Pi_n x)$ (for $\mathbb{E}F(X_n(\cdot)) = \mathbb{E}F_n(X_n(\cdot))$ and $\mathbb{E}f(X_n(T)) = \mathbb{E}f_n(X_n(T))$).

Next we have to prove that $X_n(t) \to X(t)$ in L^2. Since this approximation is important in practice, we shall give a more general result (covering, for instance, the method of finite elements in PDE).

Proposition 5.1 *Assume Hypothesis* 2.1 *and* (L). *Suppose moreover*

Hypothesis 5.1 *there exists a sequence $(H_n, A_n, F_n, B_n)_{n\in\mathbb{N}}$ where*

(i) *for each n, H_n is a finite dimensional subspace of $\mathbb{D}(A)\bigcap\mathbb{D}(F)$, and $|\Pi_n x - x| \to 0,\ \forall x \in H$, where Π_n is the orthogonal projection from H to H_n;*

(ii) *$A_n : H_n \to H_n$ is linear and bounded such that for any $x \in \mathbb{D}(A)$, there are $x_n \in H_n, n \geq 0$ such that $x_n \to x$ and $A_n x_n \to Ax$ in H;*

(iii) *$F_n : H_n \to H_n$ are uniformly Lipschitzian and $F_n(\Pi_n x) \to F(x)$ in H for all $x \in H$;*

(iv) *$B_n : H_n \to L(U \to H_n)$ is strongly continuous (i.e., $\forall u \in U$, $x \to B_n(x)u$ is continuous from H_n to H_n), and for all n, $e^{tA_n}B_n(x)$ satisfies (2.2) with the same $K(t)$ on H_n (in place of $S(t)B(x)$), and*

$$\|S_n(t)B_n(\Pi_n x) - S(t)B(x)\|_2 \to 0, \quad \forall\, t > 0,\ x \in H. \tag{5.3}$$

Let $X_n(t)$ be the unique solution to the SDE in H_n below,

$$dX_n(t) = (A_nX_n(t) + F_n(X_n(t)))dt + B_n(X_n(t))dW(t), \quad X_n(0) = x_n \in H_n, \tag{5.4}$$

where $x_n \to x$ in H, and $X(t)$ be the unique mild solution to the SDE (A, F, B). *Then*

$$\sup_{t\leq T}\mathbb{E}|X_n(t) - X(t)|^2 \to 0, \quad \forall\, T > 0. \tag{5.5}$$

Proof In the proof below A_n, F_n, B_n will be identified as $A_n\Pi_n, F_n(\Pi_n x), B_n(\Pi_n x)$ defined on the whole H, where Π_n is the orthogonal projection from H to H_n.

The strong solution $X_n(t)$ to the finite dimensional SDE (5.4) is a mild solution, i.e.,

$$X_n(t) = S_n(t)x_n + \int_0^t S_n(t-s)F_n(X_n(s))ds + \int_0^t S_n(t-s)B_n(X_n(s))dW(s),$$

where $S_n(t) = e^{tA_n}$. By [7, Theorem 5.3.1 and its proof], there is a unique L^2-mild solution $X(t)$ of the SDE (2.1), i.e., $(X(t))$ is adapted, $t \to X(t)$ is a continuous mapping from $\mathbb{R}^+$ to $L^2(\Omega, \mathcal{F}, \mathbb{P}; H)$ and

$$X(t) = S(t)x + \int_0^t S(t-s)F(X(s))ds + \int_0^t S(t-s)B(X(s))dW(s).$$

Therefore, we obtain

$$\begin{aligned}\mathbb{E}|X_n(t) - X(t)|^2 \leq 3\Big\{&|S_n(t)x_n - S(t)x|^2 \\ &+ \mathbb{E}\Big|\int_0^t [S_n(t-s)F_n(X_n(s)) - S(t-s)F(X(s))]ds\Big|^2 \\ &+ \mathbb{E}\int_0^t \|S_n(t-s)B_n(X_n(s)) - S(t-s)B(X(s))\|_2^2 ds\Big\}.\end{aligned} \tag{5.6}$$

Below $T > 0$ is fixed and all t are less or equal to T. Let us control the three terms at the right hand side (r.h.s.) of (5.6) in three points.

(1) By the Trotter-Kato theorem (see [17]) and Hypothesis 5.1(ii),

$$\sup_{t\le T}|S_n(t)x - S(t)x| \to 0, \quad \forall\, x \in H \quad \text{and} \quad \sup_{n\ge 0, t\le T}\|S_n(t)\| \le C(T) < +\infty. \tag{5.7}$$

Hence for any $\varepsilon > 0$, when n is large enough, we have

$$\sup_{t\le T}|S_n(t)x_n - S(t)x|^2 \le \sup_{t\le T}\|S_n(t)\| \cdot |x_n - x| + \sup_{t\le T}|S_n(t)x - S(t)x| < \varepsilon.$$

(2) For the second term at the r.h.s. of (5.6), notice that for all $s \le t \le T$,

$$\begin{aligned}
&|S_n(t-s)F_n(X_n(s)) - S(t-s)F(X(s))| \\
&\le |S_n(t-s)[F_n(X_n(s)) - F_n(X(s))]| + |[S_n(t-s) - S(t-s)]F_n(X(s))| \\
&\quad + |S(t-s)[F_n(X(s)) - F(X(s))]|.
\end{aligned} \tag{5.8}$$

We have by the uniform Lipschitzian condition on F_n and (5.7),

$$\begin{aligned}
|S_n(t-s)[F_n(X_n(s)) - F_n(X(s))]| &\le \sup_{t\le T}\|S_n(t)\||F_n(X_n(s)) - F_n(X(s))| \\
&\le C(T)L|X_n(s) - X(s)|,
\end{aligned}$$

where $L := \sup_n \|F_n\|_{\text{Lip}}$ and $C(T)$ is given by (5.7). And for $u = t - s \in [0, T]$,

$$|[S_n(u) - S(u)]F_n(X(s))| \le |[S_n(u) - S(u)]F(X(s))| + 2C(T)|F_n(X(s)) - F(X(s))|.$$

By [7, Theorem 5.3.1], we have the estimate below

$$\sup_{t\le T}\mathbb{E}|X(t)|^2 \le C_{1,T}(1 + |x|^2). \tag{5.9}$$

Thus by the dominated convergence theorem and (5.7),

$$\int_0^T \mathbb{E}\sup_{u\le T}|[S_n(u) - S(u)]F_n(X(s))|^2 ds \to 0.$$

For the last term in (5.8), since $|F_n(X(s))| \le |F_n(0)| + L|X(s)|$ and $F_n(x) \to F(x)$ on H, we have by the dominated convergence theorem,

$$\int_0^T \mathbb{E}\sup_{s\le t\le T}|S(t-s)[F_n(X(s)) - F(X(s))]|^2 ds \to 0.$$

Summarizing the estimates above, we have for any $\varepsilon > 0$, if n is large enough, then for all $t \in [0, T]$,

$$\begin{aligned}
&\mathbb{E}\Big|\int_0^t [S_n(t-s)F_n(X_n(s)) - S(t-s)F(X(s))]ds\Big|^2 \\
&\le T\int_0^t \mathbb{E}|S_n(t-s)F_n(X_n(s)) - S(t-s)F(X(s))|^2 ds \\
&\le \varepsilon + C_T\int_0^t \mathbb{E}|X_n(s) - X(s)|^2 ds,
\end{aligned} \tag{5.10}$$

where $C_T = T[C(T)L+1]^2$.

(3) We now go to control the third term at the r.h.s. of (5.6). By the triangular inequality we have for $u = t-s \in (0,T]$,

$$\begin{aligned}&\|S_n(u)B_n(X_n(s)) - S(u)B(X(s))\|_2\\ &\le \|S_n(u)[B_n(X_n(s)) - B_n(X(s))]\|_2 + \|[S_n(u)B_n(X(s)) - S(u)B(X(s))\|_2. \qquad (5.11)\end{aligned}$$

For the first term at the r.h.s. above, we have

$$\|S_n(u)[B_n(X_n(s)) - B_n(X(s))]\|_2 \le K(u)|X_n(s) - \Pi_n X(s)| \le K(u)|X_n(s) - X(s)|.$$

For the last term at the r.h.s. of (5.11), we have by condition (5.3) that

$$\|S_n(u)B_n(X(s)) - S(u)B(X(s))\|_2 \to 0, \quad \text{a.s.}$$

Furthermore, by Hypothesis 5.1, for any $u > 0$ and $s \ge 0$,

$$\|S_n(u)B_n(X(s)) - S(u)B(X(s))\|_2 \le 2K(u)(1 + |X(s)|).$$

Thus by (5.9), we get by the dominated convergence

$$\mathbb{E}\|S_n(u)B_n(X(s)) - S(u)B(X(s))\|_2^2 \to 0.$$

We now show that this convergence is uniform in $s \in [0,T]$. Since $s \to X(s)$ is continuous from $[0,T]$ to $L^2(\Omega, \mathbb{P}; H)$ (by the proof of Theorem 5.3.1 in [6]), and $S_n(u)B_n(\cdot) : H \to L_2(U \to H)$, $n \ge 0$ are uniformly Lipschitzian in n, we see easily that $f_n(s) := \mathbb{E}\|S_n(u)B_n(X(s)) - S(u)B(X(s))\|_2^2$, $n \ge 0$ are equi-continuous and uniformly bounded on $[0,T]$. Thus the desired uniform convergence follows by the Arzela-Ascoli theorem. That uniform convergence implies

$$\begin{aligned}&\sup_{t\le T} \int_0^t \mathbb{E}\|S_n(t-s)B_n(X(s)) - S(t-s)B(X(s))\|_2^2 ds\\ &= \sup_{t\le T} \int_0^t \mathbb{E}\|S_n(u)B_n(X(t-u)) - S(u)B(X(t-u))\|_2^2 du\\ &\le \int_0^T \sup_{s\le T} \mathbb{E}\|S_n(u)B_n(X(s)) - S(u)B(X(s))\|_2^2 du \to 0.\end{aligned}$$

Thus for any $\varepsilon > 0$, we have for all sufficiently large n,

$$\begin{aligned}&\mathbb{E}\int_0^t \|S_n(t-s)B_n(X_n(s)) - S(t-s)B(X(s))\|_2^2 ds\\ &\le \varepsilon + 2\int_0^t K^2(t-s)\mathbb{E}|X_n(s) - X(s)|^2 ds \quad \text{for all} \quad t \in [0,T]. \qquad (5.12)\end{aligned}$$

Finally, summarizing point (1), (5.10) in (2) and (5.12) above, we obtain from (5.6) that for any $\varepsilon > 0$, if n is large enough, say $n \ge N$, then for all $t \in [0,T]$,

$$\mathbb{E}|X_n(t) - X(t)|^2 \le 9\varepsilon + 2\int_0^t [C_T + 2K(t-s)^2]\mathbb{E}|X_n(s) - X(s)|^2 ds. \qquad (5.13)$$

Furthermore, (5.13) remains true if one replaces the deterministic initial conditions by $X_n(0) = \xi_n \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; H)$, $X(0) = \xi \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; H)$ such that $\mathbb{E}|\xi_n - \xi|^2 \to 0$, by exactly the same proof.

Let $0 < T_0 \le T$ be determined by

$$2\int_0^{T_0} [C_T + 2K(s)^2] ds \le \frac{1}{2},$$

which is independent of n and initial conditions (only N above may depend on them). From (5.13), we see that for $n \ge N$,

$$\frac{1}{2} \sup_{t \le T_0} \mathbb{E}|X_n(t) - X(t)|^2 \le 9\varepsilon.$$

Thus $\mathbb{E}|X_n(t) - X(t)|^2 \to 0$ uniformly over $[0, T_0]$. Note that $X_n(t+T_0), X(t+T_0)$ are mild solutions of the same SDEs with W_t substituted by $W_{t+T_0} - W_{T_0}$ and with initial conditions $X_n(T_0), X(T_0)$ which are independent of $(W_{t+T_0} - W_{T_0})_{t \ge 0}$ and satisfy $\mathbb{E}|X_n(T_0) - X(T_0)|^2 \to 0$. As (5.13) still holds for $X_n(\,\cdot + T_0), X(\,\cdot + T_0)$ in place of $X_n(\,\cdot\,), X(\,\cdot\,)$, we obtain that $\mathbb{E}|X_n(t) - X(t)|^2 \to 0$ uniformly over $[T_0, 2T_0] \bigcap [0, T]$. Repeating the argument we get $\mathbb{E}|X_n(t) - X(t)|^2 \to 0$ uniformly over $[kT_0, (k+1)T_0] \bigcap [0, T]$ for all $k \ge 0$.

The proof of the proposition is completed.

Proof of Theorem 2.1 Let A_n, F_n, B_n be given by (5.1). By Hypothesis 2.3, (A_n, F_n, B_n) satisfies Hypothesis 5.1. Thus $X_n(\,\cdot\,) \to X(\,\cdot\,)$ in $L^2(\Omega, \mathbb{P}; L^2([0,T])) = L^2(\Omega \times [0,T], \mathbb{P} \otimes dt)$, by Proposition 5.1. Now parts (a) and (b) of Theorem 2.1 except the result for μ follow from Lemma 5.1 by the general Lemma 4.1. Since $P_T(x, \cdot\,) \in T_2\big(\frac{M^2}{2\delta}\big)$, and $P_T(x, \cdot\,) \to \mu$ weakly by Remarks 2.4 as $T \to +\infty$, we obtain $\mu \in T_2\big(\frac{M^2}{2\delta}\big)$. Thus we have by [2] (their argument on $\mathbb{R}^d$ works again on H) the following Poincaré inequality,

$$\mathrm{Var}_\mu(f) \le \frac{M^2}{2\delta} \mu(|\nabla f|^2), \quad \forall\, f \in C_b^1(H).$$

This implies (2.7) in part (c).

6 Proof of Theorem 2.2: Yosida's Approximation

Lemma 6.1 *Assume that $F - \eta$ is m-dissipative on a Banach space. Then the Yosida approximations F_α of F are Lipschitz continuous. Moreover, for any $\alpha > 0$ such that $1 - \alpha\eta > 0$, $F_\alpha - \frac{\eta}{1-\alpha\eta}$ are dissipative. Here the Yosida approximation F_α of F is given by*

$$F_\alpha(x) = \frac{1}{\alpha}[(I - \alpha F)^{-1}(x) - x], \quad x \in H.$$

Proof Since $F - \eta$ is m-dissipative, its *Yosida approximations* $(F - \eta)_\beta$, $\beta > 0$ are dissipative. Now for $\beta > 0$ such that $1 + \beta\eta > 0$, we have

$$\begin{aligned}(F-\eta)_\beta(x) &= \frac{1}{\beta}[(I - \beta(F-\eta))^{-1}(x) - x] = \frac{1}{\beta}\Big[\frac{1}{1+\beta\eta}\Big(I - \frac{\beta}{1+\beta\eta}F\Big)^{-1}(x) - x\Big] \\ &= \frac{1}{\beta(1+\beta\eta)}\Big[\Big(I - \frac{\beta}{1+\beta\eta}F\Big)^{-1}(x) - x\Big] - \frac{\eta}{1+\beta\eta}x \\ &= \frac{1}{(1+\beta\eta)^2}[F_{\frac{\beta}{1+\beta\eta}}(x) - \eta(1+\beta\eta)x].\end{aligned}$$

Thus $F_{\frac{\beta}{1+\beta\eta}} - \eta(1+\beta\eta)$ is dissipative. Let $\alpha = \frac{\beta}{1+\beta\eta}$. Then $F_\alpha - \frac{\eta}{1-\alpha\eta}$ is dissipative.

Proof of Theorem 2.2 For $\alpha > 0$ such that $1 - \alpha\eta_2 > 0$, we consider the approximating SDE,

$$\begin{cases} dX_\alpha(t) = (AX_\alpha(t) + F_\alpha(X_\alpha(t)))dt + BdW(t), \\ X_\alpha(0) = x, \end{cases} \tag{6.1}$$

where $F_\alpha(x) = \frac{1}{\alpha}[(I - \alpha F)^{-1}(x) - x],\ x \in H$ is the *Yosida approximation* of F.

By Lemma 6.1, F_α are Lipschitzian and

$$\langle x - y, A(x-y) + F_\alpha(x) - F_\alpha(y)\rangle \le \Big(\eta_1 + \frac{\eta_2}{1-\alpha\eta_2}\Big)|x-y|^2.$$

Since $\delta := -(\eta_1 + \eta_2) > 0$, we have for all $\alpha > 0$ small enough,

$$\delta(\alpha) := -\Big(\eta_1 + \frac{\eta_2}{1-\alpha\eta_2}\Big) > 0, \qquad \lim_{\alpha\to 0+} \delta(\alpha) = \delta_0.$$

In such case, we have by Theorem 2.1,

$$\mathbb{P}_{\alpha,x} \in T_2\Big(\frac{\|B\|^2}{\delta(\alpha)^2}\Big),$$

where $\mathbb{P}_{\alpha,x}$ is the probability distribution of the mild solution starting from x of (6.1). Furthermore by Lemma 5.1, the approximation in the proof of Theorem 2.1 and Lemma 4.2,

$$\mathbb{P}_{\alpha,x} \in \mathrm{logS}\Big(\frac{\|B\|^2}{\delta(\alpha)^2}\Big) \quad \text{on} \quad L^2([0,T],H),$$

$$P_{\alpha,t}(x,\,\cdot\,) := \mathbb{P}(X_\alpha(t,x) \in \,\cdot\,) \in \mathrm{logS}\Big(\frac{\|B\|^2}{2\delta(\alpha)}\Big) \quad \text{on} \quad H.$$

When $x \in K$, from the proof of [6, Theorem 5.5.8], we see that as $\alpha \to 0$, with probability one,

$$X_\alpha(t) \to X(t),$$

uniformly in $t \in [0,T]$. Then by Lemmas 4.1 and 4.2, we obtain parts (a), (b) and (c) of Theorem 2.2 except the inequalities for the invariant measure μ. The inequalities for μ follow from those for $P_T(x,\,\cdot\,)$ by letting $T \to \infty$.

Finally for $x \in H$, let $(x_n)_{n\ge 0}$ be a sequence in K such that $|x_n - x| \to 0$. By the proof of [6, Theorem 5.5.8] again, the mild solution $X(\,\cdot\,, x_n)$ with initial condition $X(0,x_n) = x_n$ converges to the unique generalized solution $X(\,\cdot\,,x)$, uniformly over $[0,T]$ with probability one. Thus parts (a), (b) and (c) of Theorem 2.2 follow from the corresponding results for $X(\,\cdot\,,x_n)$ proved above, by Lemmas 4.1 and 4.2.

Finally for part (d), obviously (2.11) follows from (2.10).

References

[1] Bakry, D. and Emery, M., Diffusions hypercontractives, Séminaire de Probabilités, Lect. Notes Math., **1123**, Springer, 1985, 259–206.

[2] Bobkov, S., Gentil, I. and Ledoux, M., Hypercontractivity of Hamilton-Jacobi equations, *J. Math. Pure Appl.*, **80**, 2001, 669–696.

[3] Bobkov, S. and Götze, F., Exponential integrablity and transportation cost related to logarithmic Sobolev inequalities, *J. Funct. Anal.*, **163**, 1999, 1–28.

[4] Djellout, H., Guillin, A. and Wu, L., Transportation cost-information inequalities and applications to random dynamical systems and diffusions, *Ann. Probab.*, **32B**(3), 2004, 2702–2732.

[5] Da Prato, G., Debussche, A. and Goldys, B., Some properties of invariant measures of non-symmetric dissipative stochastic systems, *Probab. Th. Rel. Fields*, **123**, 2002, 355–380.

[6] Da Prato, G. and Zabczyk, J., Stochastic Equations in Infinite Dimensions, Encyclopedia of Mathematics and Its Applictions, Combridge University Press, 1992.

[7] Da Prato, G. and Zabczyk, J., Ergodicity for Infinite Dimensional Systems, London Mathematical Society, Lecture Note Series 229, Combridge University Press, 1996.

[8] Fang, S. and Shao, J., Transportation cost inequalities on path and loop groups, *J. Funct. Anal.*, to appear.

[9] Feyel, D. and Ustunel, A. S., Measure transport on Wiener space and Girsanov theorem, *CRAS Serie I*, **334**, 2002, 1025–1028.

[10] Feyel, D. and Ustunel, A. S., The Monge-Kantorovitch problem and Monge-Ampère equation on Wiener space, *Probab. Theor. Rel. Fields*, **128**(3), 2004, 347–385.

[11] Gourcy, M. and Wu, L., Logarithmic Sobolev inequality for diffusions w.r.t. the L^2-metric, preprint, 2004.

[12] Gross, L., Logarithmic Sobolev inequalities, *Amer. J. Math.*, **97**, 1975, 1061–1083.

[13] Ledoux, M., The Concentration of Measure Phenomenon, Mathematical Surveys and Monographs, Vol. 89, Amer. Math. Soc., 2001.

[14] Marton, K., A measure concentration inequality for contracting Markov chains, *Geom. Funct. Anal.*, **6**, 1997, 556–571.

[15] Marton, K., Bounding $\bar{d}$-distance by information divergence : a method to prove measure concentration, *Ann. Probab.*, **24**, 1996, 857–866.

[16] Otto, F. and Villani, C., Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality, *J. Funct. Anal.*, **173**, 2000, 361–400.

[17] Pazy, A., Semigroups of Linear Operayors and Applications to Partial Differential Equations, Applied Mathematical Sciences, **44**, Springer-Verlag, New York, 1983.

[18] Talagrand, M., Transportation cost for Gaussian and other product measures, *Geom. Funct. Anal.*, **6**, 1996, 587–600.

[19] Villani, C., Topics in optimal transportation, Grad. Stud. Math. Vol. 58, Amer. Math. Soc., 2003.

[20] Wang, F. Y., Transportation cost inequalities on path spaces over Riemannian manifolds, *Illinois J. Math.*, **46**, 2002, 1197–1206.

[21] Wang, F. Y., Probability distance inequalities on Riemaniann manifolds and path spaces, *J. Funct. Anal.*, **206**, 2004, 167–190.

[22] Wu, L. and Zhang, Z. L., Talagrand's T_2-transportation inequality w.r.t. a uniform metric for diffusions, *Acta Math. Appl. Sinica, English Series*, **20**(3), 2004, 357–364.